Dell CloudEdge C6145 Evaluation - 01/05/2012

Chris Hollowell <hollowec@bnl.gov>
RHIC/ATLAS Computing Facility - Brookhaven National Laboratory

A pre-production Dell CloudEdge C6145 server based on the AMD Interlagos processor was evaluated for possible future procurement at the RHIC/ATLAS Computing Facility at Brookhaven National Laboratory. The test system consisted of two independent (with the exception of a shared power supply) quad-processor motherboards in a 2U chassis. Several benchmarks were run on the machine, and compared to the performance of a Dell PowerEdge R410 server based on the Intel Westmere processor. For the purposes of our performance tests, only one of the motherboards in the C6145/Interlagos chassis was utilized.

Detailed Specifications

Component	**Dell CloudEdge C6145 (2U)**	**Dell PowerEdge R410 (1U)**
Processors	4 AMD Opteron 6276 (Interlagos) processors @2.3 GHz	2 Intel Xeon X5650 (Westmere) processors @2.67 GHz
Physical Cores	16 cores/processor = 64	6 cores/processor = 12
Logical Cores	No hyperthreading support = 64	Hyperthreading enabled = 24
RAM	128 GB, 1333 MHz DDR3	48 GB, 1333 MHz DDR3
Disk Configuration	8 500 GB SATA (3 Gbps, 7200 RPM) Drives, RAID0	4 1 TB SATA (3 Gbps, 7200 RPM) Drives, RAID0
NIC	1000baseT	1000baseT

An issue was encountered installing our Scientific Linux (SL) 5.3 OS on the C6145 host: the LSI SAS2008 disk controller was not supported by the install kernel. Therefore, we had to use SL 5.7 installation images (with a stock updated mpt2sas driver) to install the system instead. HEPSPEC06 and ATLAS software benchmarks were run on both the C6145 and R410 systems under SL5. All input/output was read/written to local disk in these tests. While the HEPSPEC06 results for the C6145 were promising, the parallel ATLAS software benchmark results were somewhat poor for this system.

HEPSPEC06 Benchmark (SL5, GCC 4.1.2)

Dell PowerEdge R410 (Westmere) – 24 parallel processes (1/logical core)	192.73
Dell CloudEdge C6145 (Interlagos) – 64 parallel processes (1/core)	472.36

ATLAS Software Benchmarks (SL5, GCC 4.3.2)*

Values are wall time to completion per average process

Single Process

Test	**Dell CloudEdge C6145 (Interlagos)**	**Dell PowerEdge R410 (Westmere)**
Setup	0:00:29	0:00:35
EvtGen	0:01:36	0:01:25
AtlasG4	0:25:40	0:16:33
Digi	0:02:24	0:02:13
Reco	0:10:22	0:10:05

24 Processes in Parallel

Test	Dell CloudEdge C6145 (Interlagos)	Dell PowerEdge R410 (Westmere)
Setup	0:25:55	0:12:42
EvtGen	0:31:13	0:10:44
AtlasG4	1:01:50	0:38:25
Digi	0:42:51	0:16:02
Reco	2:24:33	0:50:35

64 Processes in Parallel

Test	Dell CloudEdge C6145 (Interlagos)	Dell PowerEdge R410 (Westmere)
Setup	1:11:18	N/A
EvtGen	2:02:17	N/A
AtlasG4	4:03:30	N/A
Digi	2:44:26	N/A
Reco	5:54:57	N/A

In particular, note that when 24 simultaneous jobs were run on each host, reconstruction took ~1.5 hours longer per process on the C6145 server than on the R410 system. When 64 reconstruction jobs (one per core) were run on the C6145 host, completion time per process increased by a factor of 6 when compared to the result of the 24-job (also one per logical core) R410 test. While the C6145 has ~2.5 more cores than the R410, the core count increase is not enough to make up for the decrease in performance per core. A recommended BIOS upgrade on the C6145 did not improve the performance of the ATLAS software on this system.

After reporting the performance issues encountered with the C6145 to AMD, they suggested that our tests be rerun under SL6, since this OS ships with a newer Linux kernel (2.6.32-based vs SL5's 2.6.18-based kernel) which was speculated to contain improvements in the NUMA memory allocator and CPU affinity code. They also recommended that the test code be recompiled with GCC 4.6.x, which has support for the enhanced AMD Bulldozer/Interlagos instruction set (XOP and FMA4 SIMD instructions). To generate binaries using these instructions, the ”-march=bdver1” option needs to be passed to the GCC compiler. However, this also makes the resulting binaries non-portable to older and non-AMD X86-64 processors.

At the time of writing, the ATLAS software cannot be successfully compiled with GCC 4.6.x. Furthermore, there is currently no production release of the ATLAS software for SL6. Therefore, we opted to run vanilla ROOT benchmarks compiled with GCC 4.6.2 on the test machines. SL 6.1 was installed on the C6145, while SL 5.3 remained on the R410. All I/O was performed to /dev/shm.

ROOT Benchmarks (GCC 4.6.2)*

Values are ROOTMARKS per average process

Test	Dell CloudEdge C6145 (Interlagos) - SL6	Dell CloudEdge C6145 (Interlagos) - -march=bdver1 SL6	Dell PowerEdge R410 (Westmere) - SL5
Single process	1484.35	1473.48	1183.43
24 Jobs	N/A	N/A	927.35
64 Jobs	405.62	444.29	661.74

While the single process ROOT benchmark results were better for the C6145/Interlagos, the parallel performance for 64 simultaneous jobs was ~40% worse for this system when compared to the R410/Westmere. AMD indicated that manually tweaking the NUMA settings per process with the 'numactl' utility could also improve performance. However, this is a somewhat complicated task, and it is currently unreasonable for us to request that our user community and/or batch system adjust the NUMA policy for interactive and batch jobs.

Conclusion

The Dell CloudEdge C6145 server, based on the AMD Interlagos processor, does not appear to be particularly well suited for use in HEP/NP HPC environments. While the system's HEPSPEC06 rating is high, parallel runs of actual HEP/NP and ATLAS code on the system performed poorly. The discrepancy between the HEPSPEC06 benchmark results, and the performance of actual ATLAS analysis/simulation/reconstruction code should be investigated further. A possible explanation may be that all/most of the data manipulated by HEPSPEC06 was able to fit into the L1/L2/L3 caches of the Interlagos processors (resulting in fairly little parallel memory access), while this was not the case for the ATLAS software.

** Thanks to Shuwei Ye <yesw@bnl.gov> for running the ATLAS software and ROOT benchmarks.*